# Statement of Issues and Justification:

# Prerequisite Criteria:

## A. How is the NRSP consistent with the mission?

**Background.** The National Animal Genome Research Program (NAGRP), NRSP-8, has been hugely successful, exceeding all expectations, by delivering complete genome sequences of seven agricultural animal species (pig, cattle, sheep, goat, horse, chicken and turkey) and providing genetic tools and resources that have revolutionized the animal breeding industry. Genome-enabled technologies developed under NRSP-8 are now integral components of commercial animal breeding and production for many species, resulting in a multifold return on investment to US stakeholders and producers. These successes, along with concomitant advances in genomics-enabled technologies, resulted in the release of the 2018-2027 USDA Animal Genome Blueprint (1) which outlines key areas for future research and funding. A crucial element identified in this report is "Science to Practice", which is the application of genomics-enabled technologies to traits and phenotypes critical to animal industries. The "Science to Practice" goals are undoubtedly attainable, in large part because of the genetic resources and technologies developed through NRSP-8.

This proposal leverages the significant accomplishments and products of the NRSP-8 program, but in contrast to previous requests for renewing NRSP-8, the new project will redirect its objectives and focus solely on capacity development. Importantly, we will expand the NRSP-8 community to include direct involvement of additional stakeholders, including non-genomics scientists and researchers, Extension personnel, and animal industry representatives. Moreover, increasing capacity that enables the application of genomics to animal traits and phenotypes will require the development of linkages with data scientists, a group that has not to date been a significant part of NRSP-8. Integration of genomics and data science, and more specifically, bioinformatics, which is the computation and analysis of biological data, will link genome data with phenome data in a very deliberate way. Animal industries are increasingly employing data collection technologies in all aspects of production, performance, health and welfare. Examples of applying data science to animal industries include image analysis to investigate tail-biting in pigs (2), Radio Frequency Identification (RFID) for studying cattle grazing behaviors (3), gait analysis of horses using neural network analysis of video data (4) and machine learning algorithms to predict health issues in aquaculture systems (5). Linking these incredibly rich datasets with genomic information and tools will provide new opportunities for US animal agriculture and accomplish the goals of the USDA Animal Genome Blueprint.

The realization of this vision will require the development of enabling capacity that is not covered under current competitive grant programs. In addition, initial conversations that have occurred between NRSP-8 constituents and industry representatives about their industry's specific needs must be continued and expanded, thereby developing knowledge, trust and understanding among the two groups. Key to these conversations are our land grant Extension personnel who are uniquely positioned to deliver new knowledge and applications to animal industries and in the opposite direction, they can ensure that current production issues are at the forefront of research objectives. In addition, novel collaborations of NRSP-8 researchers with a broad range of data scientists will further develop capacity for integrating new data types into animal agriculture research. Underpinning these efforts, we must ensure that a broad and diverse group of animal

researchers, beyond just animal genomicists, are prepared to utilize data science techniques (including bioinformatics) to support innovation in US animal industries. With this in mind, *we propose a national multi-institutional capacity project to develop the infrastructure and expertise required to apply genomics-enabled technologies to US animal industries.*

**Supporting the NRSP Mission.** The overall goal and specific objectives of this proposal are directly aligned with the mission of the NRSP through development of enabling technologies and by providing training and education that support the application of genomics across the complete range of agricultural animal sciences and species. By bringing together researchers in animal science, data science and genomics, as well as Extension faculty and industry representatives, we will accomplish what could not be realized by individual efforts or by animal genomicists alone. The opportunity to coordinate discussions and develop collaborations will provide a set of guiding principles and resources which can then be applied to research projects across the country. Moreover, we anticipate that the opportunities provided by multi-disciplinary interactions will create novel Research, Education and Extension links focused around the application of genomics to animal industries.

**Supporting State Agricultural Experiment Stations:** The Land Grant mission is focused on student-centered education to develop the 21st century workforce, delivery of cutting-edge discoveries that advance knowledge in state and national need areas, and providing outreach that educates and elevates individuals, families and communities. This proposal, which focuses on the development of capacity, responds to all three areas. In addition, the application of genomics-enabled technologies ensures that US agriculture remains globally competitive, enhances US food security and safety, supports sustainable production innovation, and provides actionable information that informs regulatory policies.

## B. How does this NRSP pertain as a national issue? (10,000 characters)

### 1. National Scope

Livestock and poultry are top US agricultural commodities, accounting for more than half of the agricultural cash receipts and totally at least $100 billion each year. Conversely, the US imports ~90% of its seafood, at a deficit of more than $17 billion per year. Therefore, advances in livestock, poultry and aquaculture production capacity and efficiency are essential for advancing national and global food security, especially with increasing consumer demands for improved animal welfare and reduced chemical interventions, the world's changing climate, and competing allocations of land and water. To enable agri-animal industries to increase production and meet growing demands, we need robust animals with superior health and production traits and optimal management of these animals, while relying to a lesser degree on antimicrobials that can increase the risk of resistant pathogens, especially emergent zoonotic pathogens (6). These challenges cannot be addressed without significant advances in the animal sciences, including nutrition, physiology (reproduction, lactation, growth, ethology, etc.), genetics, and meat science. Scientific efforts must capitalize on the latest advances in technology, including the various fast-developing "omics technologies'' (genomics, epigenomics, transcriptomics, proteomics, metabolomics, microbiomics), automated high-throughput phenotyping technologies (sensors, cameras, etc.), and the associated statistical "big data" approaches (e.g., artificial intelligence and machine learning). However, application of these emerging technologies puts additional pressures on the broader animal science community, including an understand of the various omics technologies and the ability to handle big data. Scientists in animal science disciplines outside of genetics are typically not trained in the use of big data, nor in the various genomics technologies. This results

in an opportunity cost, because animal scientists may not fully exploit the freely available genomic tools, resources, and knowledge that would greatly benefit and illuminate their research. Thus, there is a need to enable the use of genomic information across all fields of animal science in both academia and industry. Also, it must be noted that funds available for generating large datasets relevant to animal genomics are extremely limited. Therefore, the re-use and repurposing of existing datasets for gaining insight into novel questions represent opportunities to increase the value of data collection both in time and resources. In short, to fully leverage investments in genomic information, we need to help *"normalize" the use of genomic information and associated technologies* and lower the barrier to entry for research groups that are less familiar with genomics resources and tools.

## 2. Continued national need for animal genomics capacity

In the past three years, there have been several developments that impact the future of animal genomics and its application to livestock production, including release of the *USDA Blueprint for Animal Genome Research*, followed by the 2020 release of the *USDA Agricultural Innovation Agenda* that signaled the intent of USDA to increase US agricultural production by 40 percent, while cutting its environmental footprint in half by 2050. The 2018 Farm Bill provided funding directed to the *Agricultural Genome to Phenome Initiative* ([AG2PI](AG2PI)), which also started in 2020. AG2PI is joint plant and animal effort to prepare research communities to embark on a large-scale effort to link genomes to biological function (phenomes) across crops and animals of importance to the agriculture sector of the US. Also in 2020, the EO 13921 *Promoting American Seafood Competitiveness and Economic Growth* detailed the need for improving American competitiveness in the global seafood market. The US animal genomics research community, specifically NRSP-8, is an integral partner to AG2PI and is ideally positioned to work collaboratively and synergistically, alongside the crop genomics community, to enable the linkage of genomics and predictive phenomics with other available biotechnologies. Supporting capacity development will ensure that investments in genomics are leveraged to generate the best societal and environmental benefits to help USDA meet its goals for US agriculture. *This renewal application builds on successes of NRSP-8 and from there, provides a logical, sustainable progression and expansion of capacity.*

We propose a national project that develops capability in using genomics-enabled technologies within diverse animal industries. In this context we define genomics-enabled technologies as any high-throughput platform that relies on genomics-wide data analyses (including transcriptomics, sequencing, proteomics, etc.), global analysis of metabolites, and the microbiome, as well as phenomics and large-scale genetics studies (including GWAS, haplotype analysis, and use of genetic markers, traits, or phenotypes). Developing this very broad capacity requires an infrastructure that supports consistent collection and use of genomics data types, their integration with rich phenotypic (meta) data, and the development of human expertise in bioinformatics and related data science techniques. Building both informatics and human capacity will ensure the application of these techniques into predictive biology that supports resilient agricultural systems.

Our overall goal is to develop the infrastructure and expertise required to apply genomics-enabled technologies to US animal industries. This overall goal is supported by three specific objectives:

**(1) Extending genomics capacity to a broader range of Animal Science stakeholders.** New genomic technologies enhance the collection of expanded "omic" data types and increased data volume. In fact, for most researchers, it is now easier to generate genomic data than it is to manage and analyze the resulting data. Aim 1 specifically addresses the

ways researchers can acquire, analyze, share and re-use genomics data types for their own programs.

(2) **Supporting capacity to integrate genomic and biological data.** New engineering technologies are now allowing animal scientists to collect biological data on a wide variety of animal traits and phenotypes. Similar to Aim 1, we now need to develop the ability to manage, analyze and integrate these large and complex biological datasets and then connect them to the genomics information that is also being collected. Aim 2 specifically addresses the need for scientists to integrate diverse data types, both biological measurements and genomics, for a more complete understanding of complex agricultural systems.

(3) **Education, training and outreach to develop a data-savvy workforce.** The animal genomics community has been relatively successful at providing bioinformatics training that supports genomics studies. However, there is an acute need to extend this capability to the entire animal science community and to ensure that animal scientists can manage and apply the expanding range of data types which are used in animal industries (e.g., genetic, genomic, epigenetic, GIS, images and audio data types). Aim 3 specifically addresses the need for ensuring that we are developing the 21st century agricultural workforce in the areas of informatics science, which can keep US agriculture globally competitive and resilient.

# Rationale:

## A. Priority Established by ESS:

This proposal supports Grand Challenges identified by the ESCOP Science and Technology Committee (STC) Science Roadmap for Food and Agriculture (2020) as outlined below.

*Grand Challenge 1: Enhancing the sustainability, competitiveness, and profitability of US food and agricultural systems.* Developing capacity that links genomics technologies with key production, performance and welfare phenotypes will address this grand challenge by:

- maximizing capacity to enhance animal production and performance using genomics and developing new animal breeds and stocks to support diverse and resilient agricultural systems;
- applying technologies to improve animal health, well-being, and welfare in all production systems;
- enhancing nutrition efficiency and sustainability, productivity, and quality of food products in agricultural systems; and
- improving technologies for animal waste utilization and management to reduce the environmental impact of agricultural production systems.

Moreover, these outcomes also directly align with USDA Strategic Plan (7) Strategic Goal 2 (*Ensure America's Agricultural System is Equitable, Resilient, and Prosperous*), particularly by protecting animal health and fostering agricultural innovation.

*Grand Challenge 2: Adapting to and mitigating the impacts of climate change on food, feed, fiber, and fuel systems in the United States.* Developing capacity to apply omics technologies to animal industries supports this grand challenge by:

- applying precision agriculture for developing resilient animal lines and breeds adapted to local and to changing conditions;
- developing and sharing new, rapid breeding technologies to effectively respond to emergent vulnerabilities as microclimates become suitable for previously non-threatening diseases and pests and freshwater resources become limited;
- generating new livestock models focused on heat stress and greenhouse gas mitigation in livestock facilities; and
- acquiring and sharing accessible phenotypic data to support decision systems that integrate animal management with changing climate forecasts.

These outcomes also directly align with USDA Strategic Plan (7) Strategic Goal 1 (*Combat Climate Change to Support America's Working Lands, Natural Resources*) by building capacity to adapt to the consequences of climate change and reducing greenhouse emissions.

*Grand Challenge 3: Supporting energy security and the development of the bioeconomy from renewable natural resources in the US.* The integration of animal genetics with phenotypes plays an important role in the US bioeconomy (8). Examples of how the capacity developed under this proposal will support this grand challenge include:

- continuing development of sustainable animal-sourced food products;
- developing animal-based bioeconomies to support the revitalization of rural areas;
- identifying and preserving biodiversity of animal-based bioeconomies; and
- developing and sharing information to support innovative systems for reducing, recycling or reusing agricultural waste.

*Grand Challenge 4: Ensuring a safe, secure, and abundant food supply for the US and the world.* Developing capacity to apply omics technologies to animal industries supports this grand challenge via:

- increased information about optimal management strategies for reducing bioactive compounds such as antibiotics and pharmaceuticals;
- information that informs effective food production regulatory policies by the USDA, FDA, Environmental Protection Agency (EPA), and other federal agencies;
- support for technologies that maximize the genomic potential of animals for enhanced productivity and quality;
- creation of novel breeding programs that balance and optimized nutritional value with production characteristics;
- development of direct collaborative links between research institutions and industry to promote the translation of new knowledge into practical applications; and
- mechanisms for cooperative international initiatives that globally enhance food safety, security, and abundance.

In addition, these outcomes support USDA Strategic Plan (7) Strategic Goal 4 (*Provide All Americans Safe, Nutritious Food*), in particular preventing foodborne illness and protecting public health.

*Grand Challenge 5: Improving human health, nutrition, and wellness of the US population.* Developing capacity for omics technologies combined with the ability to capture key industry phenotypes will support this challenge in the following ways:

- develop animal food products with enhanced nutrition;

- enhance the use of animal models in human medicine; and
- support the use of animal-assisted therapies in human rehabilitation and wellness.

*Grand Challenge 6: Heightening environmental stewardship through the development of sustainable management practices.* The capacity developed by this proposal underpins innovations that can support:

- sustainable feeding and pest management strategies for livestock production systems;
- mitigation of methane emission from ruminants through combined genetic-dietary approaches;
- enhanced feeding practices to reduce nitrogen waste from livestock systems;
- support for precision agriculture approached to reduce chemical/antibiotic use and waste runoff;
- information to increase our understanding of ecological interactions that occur in animal industries; and
- advanced production of fish, shellfish, and aquatic plants in aquaculture systems through application of omics technologies and advanced selective breeding and domestication.

In addition to its alignment with the ESCOP STC Grand Challenges and the USDA Strategic Plan (2022-2026), the objectives of this proposal align directly with the USDA Blueprint for Animal Genome Research (1). This report highlighted progress towards assembling genomes for agriculturally relevant animals and identifying genomic and sequence variants. The report also provided examples of how these techniques and knowledge had been applied to animal industries. However, the report notes that "understanding these genomic effects is now limited by the phenotypes that are collected". Finally, the 2021 Threats to Food and Agricultural Resources report released by the US DHS and ODNI Analytic Exchange Program outlined grand challenges to US food security (9) including: "*The US government should lead research coordination of public-private partnerships for [agricultural] information sharing standards and risk mitigation*" and "*The US government needs to promote domestic aquaculture for food production*".

In sum, the goal of this proposal is to make genomic techniques and knowledge widely accessible to a broad range of researchers and stakeholders and to leverage data science expertise so that phenomic data collection and integration with genomic information are expanded, enhanced and coordinated. These actions will significantly increase national food animal production capabilities.

## B. Relevance to stakeholders:

### 1. Stakeholders and their needs

Stakeholders will be included in this project based on their ability and commitment to enhancing the application of genomics into the animal industries. These groups are:

- <u>Animal science researchers:</u> This proposal supports the development of genomics expertise that will lower barriers for a broad group of researchers so they can effectively apply genomics to their research.
- <u>Future animal scientists:</u> This proposal will provide education and training opportunities for animal science students, as well as develop links with industry stakeholders who can provide internship placements and support novel collaborations for the students.
- <u>Breed associations:</u> An important goal of this proposal is to engage with representatives of breed associations to identify their needs and to ensure that they can access and apply resources, tools and expertise developed by this proposal.

- Animal breeders: We will continue to support the needs of animal breeders by developing shared cyberinfrastructure and providing training opportunities.
- Bioinformatic and data scientists: Many bioinformatics trained personnel are already involved in the NRSP-8 project. By including a broader group of informaticians, we will support the development of data science capacity specifically for animal agriculture and engage with data scientists who are interested in working on agricultural systems but to date have not been involved.
- Regulatory and policy offices: This project will develop capacity which helps federal agencies to apply evidence-based science and informs decision making for new policies. Well-designed and achievable polices and regulations are critical components in the application of genomics to animal industries.
- USDA directors and managers: Discussions with all stakeholders are expected to identify emerging industry needs. We will continue to work closely with USDA administration, including AFRI program managers, to develop RFPs and white papers that advance research in the needed areas.

As the capacity for applying genomics-enabled technologies to US animal industries increases, we anticipate that additional stakeholder groups will be added to the project. Furthermore, we acknowledge that the ultimate stakeholders of this NRSP are consumers and US taxpayers. Members of this project will strive to engage diverse sectors and communities in order to increase awareness and knowledge of genomics and its contribution in ensuring the sustainability of US animal agriculture.

**Involving stakeholders in this proposal:** As detailed in the business plan, stakeholders will be engaged in multiple aspects of this project, including key decision-making processes. First, we will establish an External Advisory Board (EAB) of industry stakeholders to ensure that project activities are focused on industry needs, and the NRSP membership will meet at least annually with the EAB. Second, we will use the existing AnGenMap and other professional mailing lists to encourage broad participation by all stakeholders, including animal scientists, breeders, Extension personnel, and educators. Third, many of the bioinformaticians involved in NRSP-8 have contacts in data science, including public data resources, databases and cyber-infrastructure platforms, as well as through professional associations. These relationships will be leveraged into this NRSP, as well as to the broader data science community, through targeted workshops and training events. This last approach will have the additional benefit of increasing awareness of opportunities for data scientists within animal agriculture. Details of how the stakeholders will be involved in this NRSP are further elaborated in the Business Plan and Integration sections.

**Assessing stakeholder use of outcomes:** Our assessment of stakeholder engagement with activities and resources will include quantitative and qualitative pre- and post- measures. Examples of assessment metrics include:

- Attendance of stakeholders at NRSP events and workshops
- Event and workshop satisfaction assessments
- Responses to project surveys on developing capacity
- Engagement on the project website and mailing list
- Resources developed by members of this project and their access/use metrics
- Research products that cite the use of resources and expertise developed under this project
- Collaborative links between NRSP members and with non-members

- Students attending education events and their assessment of learning goals for these events
- Internship and training opportunities supported by NRSP members and industry stakeholders.

Assessment of NSRP outcomes will be included as part of each activity. The outcomes will be reviewed annually by project members and stakeholders to ensure we continue to meet our objectives and effectively engage stakeholder groups. We anticipate that review of project outcomes and stakeholder involvement will require us to adapt to changing circumstances and emerging research objectives as the project progresses.

**Contribution to public policy:** Project outcomes will include resources and information that support the application of genomics to animal industries. Making genomics data more accessible and the associated analyses more reproducible will aid in development of or contribute to the discussion of public policy related to food safety, environmental protection and understanding the impacts of genetically modified animals. By increasing capacity that enables a wider range of agricultural scientists to apply genomic technologies, we broaden participation and scientific expertise that address key challenges in animal industries such as reducing the use of antibiotics in production systems, identifying approaches for animal waste utilization, improved animal health and welfare, and understanding the mechanisms and effects of genetic modifications both on animals and their production environment. US regulatory agencies will be able to utilize detailed information about the molecular basis of complex biological systems, which supports the application of evidence-based science for policy development. We will include representatives from key regulatory agencies in NRSP discussions and activities so that a diverse group of stakeholder voices are incorporated during the early stages of the project as well as in the planning of subsequent resources and directions.

## 2. Renewal Justification

This capacity proposal directly supports and extends the genomics capacity developed under NRSP-8, which currently has 105 members from 49 institutes. Members span 36 US states and seven countries, making it truly a national project with international collaborators. From 2018-2021, NRSP-8 was supported by OTT Multi-state Research Funding at $500,000 per year. During that same 4-year period, NRSP-8 members produced 924 publications and obtained over $43 million in competitive funding - a return on investment of more than $20 for every dollar provided to the NRSP-8 project. With a focus on future impacts, the proposed NRSP project will accelerate use and application of genomics resources that have been developed under NRSP-8, and this will be accomplished through training and education that supports assimilation of genomics-technologies across all animal science areas.

Over the last four years, NRSP-8 members have held meetings with industry stakeholders to gauge the application of genomics in their companies and production systems. Industry representatives have made it clear they want continued access to bioinformatics resources, expanded bioinformatics training, opportunities to recruit data-savvy graduates, and the ability to link genetics data with information they are collecting on animal traits and phenotypes within their operations. Thus, this new NRSP proposal will have a concentrated emphasis on animal industry needs. In demonstration of this emphasis, letters of support from industry stakeholders are included in the proposal.

# Implementation:

## A. Objectives and Projected Outcomes:

The aims of this proposal are focused on developing informatics capacity to meet a variety of stakeholder needs (see previous sections). Also, accomplishments and productivity of the previous NRSP-8 project are discussed in the previous section.

**Specific Aim 1: Extending genomics capacity to a broader range of Animal Science stakeholders.**

**Approach:**

- Survey stakeholders on data capacity needs
- Design workflows to simplify data sharing
- Develop resources (e.g., protocols.io) for genomic workflows
- Support community use of micropublications that ensure datasets are publicly available and citable
- Generate community standards for data reuse

**Specific Aim 2: Supporting capacity to integrate genomic and biological data.**

**Approach:**

- Workshops to identify stakeholder data collection and data analysis gaps
- Roundtable discussions with data scientists to develop interdisciplinary teams
- Support for students and early career investigators, including attendance at bioinformatics and data science meetings

**Specific Aim 3: Education, training and outreach to develop a data-savvy workforce.**

**Approach:**

- Development and publication of stand-alone educational modules that can be added to existing animal science courses or used as stand-alone training
- Linking educational modules with micro-certificates that signal data competencies
- Development of a mailing list/bulletin board to publicize internship opportunities between stakeholders, students and academic advisors in animal science
- Bioinformatic training workshops at meetings and conferences

## B. Expected Outcomes and Impacts

**Outputs:**

Published report on the data management capacity needs of stakeholders. Expected impact: Knowledge to support the application of precision genomics in animal industries.

Coordination of resources to facilitate data deposition and re-use. Expected impact: Support for data sharing and re-use to accelerate the use of omics technologies and ensure a return on investment.

Evaluation and standardization of methods for recognition of data reuse. Expected impact: Capacity for the broader research community to effectively utilize genomics approaches.

Exchange of ideas to identify key data collection and management needs of stakeholders. Expected impact: Development of knowledge, resources and interdisciplinary teams that can address the Grand Challenges facing agriculture.

Published report of animal industry needs for integrating genomics and other data types. Expected impact: Capacity for precision agriculture in animal industries.

Identification of future stakeholder needs integrating genomics data. Expected impact: Support for the development of sustainable, resilient and economically viable animal industries in a changing environment.

Exchange of ideas and expertise with data science experts. Expected impact: New collaborations between scientists of diverse research areas.

Support for students and early career investigators to attend bioinformatics and data science meetings. Expected impact: Development of a data-literate animal agriculture workforce.

Coordination of undergraduate and graduate student educational resources. Expected impact: Development of a data-literate animal agriculture workforce.

Identification and coordination of information about internship programs for animal science students. Expected impact: Preparation of a new generation of data-savvy scientists for the agricultural workforce.

Coordination to develop animal science relevant training. Expected impact: Opportunities for continued professional development for the animal industry workforce and partnerships between industry and academia.


# B. Management, Budget, and Business Plan: (16,000 characters)

## 1. Business Plan:

**Rationale for Modest and Sustained Support:** In the last renewal period, NRSP-8 received $500,000 per year from SAES which was distributed across six species coordinators (pig, cattle, sheep/goat, horse, poultry and aquaculture at $65K per coordinator) and a bioinformatics coordinator ($110K). Members of the NRSP-8 project have been highly productive, attributing 924 publications and more than $43 million in competitive grant and contract funding to NRSP-8 from 2018 to 2021. However, the broader and more lasting impacts of the NRSP-8 project are the genomics resources and tools that were developed since the inception of NRSP-8 in 1993. These tremendous outputs will now be leveraged to future genomic discoveries in farm animal species under the new proposed project, while expanding capacity to a much broader stakeholder

community. The new NRSP project's business management plan is designed to ensure that the project will function smoothly and effectively and will maximize opportunities to seek funding beyond the OTT MRF. Modest annual funding ($127,120) is requested for the first three years of the project.

**Organizational Structure:** An outline of the organizational structure for this project is shown in Figure 1 and indicates interaction among stakeholders, NRSP leadership and NRSP members.
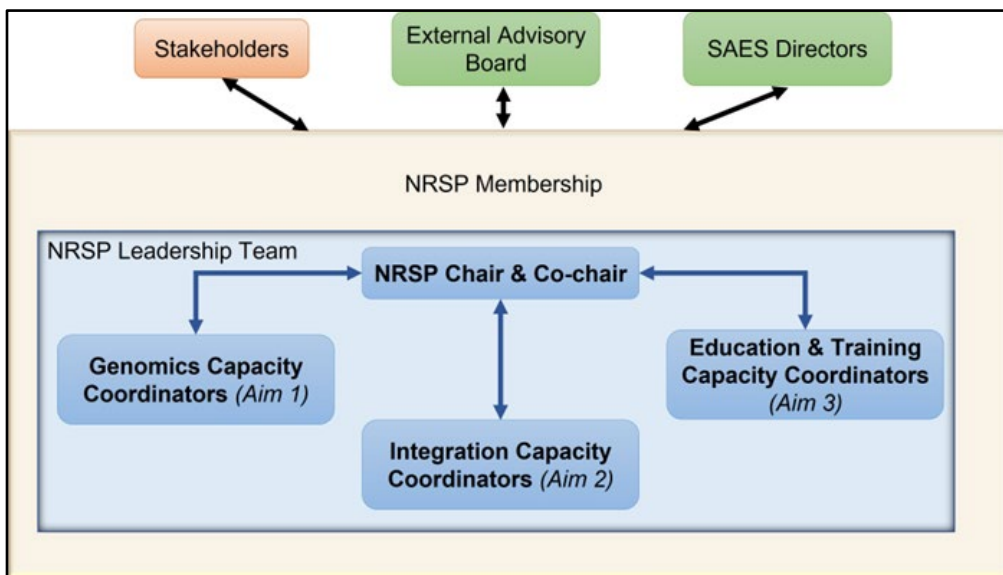


**Figure 1. Organizational structure for NRSP project management.**

Briefly, the NRSP Leadership Team will consist of eight elected positions: Chair, Co-chair, and Coordinators and Co-Coordinators for each of the three specific aims described above. The leadership team will report to SAES Directors, stakeholders, and an External Advisory Board (EAB).

NRSP Leadership Team:

- Chair (elected by NRSP members): manage the logistics of the project and provide annual reports to the membership, stakeholders, SAES Directors, and EAB.
- Co-chair (elected by NRSP members): assist Chair with project responsibilities and serve as the Chair-Elect. The co-chair will move into the position of Chair after the Chair's term is completed, and a new Co-Chair will be elected by the NRSP membership.
- Coordinators (elected by NRSP members): manage activities for their respective specific aim and report on these activities and their outcomes at the annual meetings.
- Co-Coordinator (elected by NRSP members): assist with Coordinator responsibilities and ensure activities meet stakeholder needs. Co-Coordinators will ideally be elected from Extension, industry, or the stakeholder group.

External Advisory Board (EAB):

The EAB will consist of five stakeholder members. The role of the EAB is to attend the annual meeting of the NRSP project, review the annual progress reports and provide the NRSP Leadership Team with feedback about progress, activities, stakeholder needs, and the impact of capacity developed within the project.

The Leadership Team and EAB positions will have three-year terms, with the possibility of a second consecutive term if re-nominated and re-elected. Nominations for membership on the NRSP Leadership Team and EAB will be sought to reflect the range of stakeholder interests, technical expertise, career stage and institutions that encompass the NRSP membership. Nominations will then be presented in an election available to all NRSP members. The first election will be held following the notification of project approval; Drs. Fiona McCarthy (University of Arizona) and Stephanie McKay (University of Vermont) will serve as interim Co-chairs of the project until that election is held.

While the budget outlines a proposed distribution of funds, there may be revisions to the budget based on discussions between the NIFA Director, the NRSP Leadership Team and the EAB in an effort to better align resources across the project's aims and activities. Thus, adjustment(s) of the budget may be requested during the project period, and approval of the revision(s) will be sought from the administrative coordinators assigned to the project.

**Planned Collaborations & Outreach:** A key aspect of this proposal are the collaborations that will be facilitated in several ways with animal genomicists and:

- researchers in other subspecialties of animal sciences and their allied fields (e.g., nutrition, reproductive biology, veterinary medicine) though workshops and meetings at existing conferences (e.g., American Society of Animal Science and associated regional conferences);
- those in other data science fields in order to develop capacity in new informatic techniques that can be applied to animal agriculture;
- groups working on open data sources that support better data management and standards within animal sciences (e.g. the FAIR Initiative);
- genomics/bioinformatics resources such as CyVerse and Galaxy to ensure support for sustainable training resources;
- collaborative links with existing projects such as the USDA-funded AG2PI, the AgBioData Consortium, and NRSP-10 National Database Resources for Crop Genomics, Genetics and Breeding Research;
- related multi-state projects that have genetic improvement/data sensing themes, including:
  - NC1170 Advanced Technologies for the Genetic Improvement of Poultry
  - NECC1901 Integrating Genomics and Breeding for Improved Aquaculture Production of Molluscan Shellfish
  - WERA1 Coordinating Beef Cattle Breeding Research and Education Programs for the Western States
  - S1086 Enhancing sustainability of beef cattle production in Southern and Central US through genetic improvement
  - S1069 Research and Extension for Unmanned Aircraft Systems (UAS) Applications in U.S. Agriculture and Natural Resources

    The integration of activities with these projects is described in more detail in other sections of the proposal.

Outreach effort will focus on:

Expanding genomics capacity for all animal scientists. NRSP annual meetings will alternate between the International Plant and Animal Genome (PAG) and other meetings routinely attended by animal scientists (e.g., the annual American Society of Animal Sciences meeting). Outreach

to other sectors of the broader community will occur through presentations, informational sessions and training opportunities at meetings attended by animal sciences researchers and other stakeholders (e.g., the Havemeyer meeting, Beef Improvement Federation Symposium, American Sheep Industry Convention). The outreach events will include presentations on how NRSP tools and resources have been applied in various research projects and in various species/industries, discussions about specific community needs and gaps, and hands-on training for non-genomicists. These outreach efforts will inform the broader community about research capacity that is either available or being developed and will include feedback so that new developments are aligned with stakeholder needs.

Developing informatics capacity for animal industries. Bioinformatics personnel routinely attend and present at informatics and association meetings (e.g., Data Carpentries, GMOD, Galaxy, CyVerse and the International Society of Biocuration). Therefore, we will use these meetings as conduits for developing cross-collaboration information exchanges with the broader informatics field. Capacity for animal scientists will be expanded through these exchanges and scientists working in informatics and data science will learn of new opportunities for applying informatics techniques to animal industries. Partnering with these informatic associations will also provide novel opportunities to develop education and training in targeted areas as well as the recruitment of additional expertise and resources for the animal science areas.

Supporting education and training capacity. Outreach to both animal scientists and informaticians will be done through education and training programs, with a primary focus on graduate and undergraduate students at land grant colleges and universities, including the 1862, 1890 and 1994 institutions. However, online educational resources will also be available to all participants of the NRSP project. Additionally, animal breeders, industry stakeholders, Extension personnel and the general public will be provided with non-technical information on genomics-enabled technologies and their application in agriculture. In this way, the successes of genomic-based research in farm animals will be disseminated and celebrated.

The project leadership will also engage with regulatory and funding agencies to ensure broad perspectives on the needs of genomics-enabled technologies across all animal industries. These conversations will also lead to funding for additional activities of this project and new research that may be developed because of capacity expansion that are outcomes of the project.

**Outcome Delivery Schedule:** The key milestones of this NRSP are described below. The project Chairs and Coordinators will be responsible for ensuring that all members and stakeholders are positioned to meet the targeted deliverables.

Year 1. *Aim 1:* Produce surveys for animal science researchers; community discussion regarding data deposition guidelines. *Aim 2:* Develop initial stakeholder and data science groups for discussions via open expressions of interest. *Aim 3:* Develop test-cases for educational modules; collect information about internship opportunities; determine needs/interest for training programs.

Year 2. *Aim 1:* Distribute survey results; develop test-cases for data deposition; organize community workshops on data reuse. *Aim 2:* Hold industry-specific round tables; develop links with data science experts. *Aim 3:* Assess training modules; review and assess internship co-ordination; deliver training workshops.

Year 3. *Aim 1:* Publish survey results; assess development of data deposition guidelines; organize follow-up with stakeholders on data reuse workshops. *Aim 2:* Complete industry-specific round tables and report on industry needs. *Aim 3:* Expand educational modules; offer and evaluate

internship experiences; expand training workshops; provide mid-project review and assessment of project activities.

Year 4. *Aim 1:* Distribute data deposition resources to the community. *Aim 2:* Host interdisciplinary discussions to foster collaborations. *Aim 3:* Add educational modules; offer and evaluate internship experiences; provide training workshops.

Year 5. *Aim 1:* Publish report on data reuse. *Aim 2:* Deliver outreach at bioinformatics, bio-curation and data science meetings. *Aim 3:* Assess education modules and revise as needed; review and assess internship coordination; assess training workshops and revise as needed; provide final report and assessment of project activities.

**Additional sources of funding:** In addition to the OTT Multi-state Research Funding (MRF) provided by SAES, members of the NRSP project will be proactive in seeking and obtaining funding from additional sources to expand resources and address issues relevant to the project. For example, NRSP-8 members have submitted a conference proposal to AFRI supporting early career scientists who want to use genomics in their research projects to attend the Advances in Genome Biology and Technology – Agriculture (AGBT-Ag) meeting in 2023. Another example of additional funding are internship opportunities that are offered by several of our industry partners. We intend to redirect these internships to this project by training students in genomics techniques and then placing the students with relevant industry partners. Additional research experiences for students will be expanded with new stakeholders offering internships. Other sources of funding that will be sought for this project include grants and industry sponsorship for meetings, and competitive funding from agencies such as AFRI, NSF and FFAR to support workforce training focused on genomics capacity.

**Sustainability:** This NRSP proposal includes provisions for sustaining capacity beyond the initial OTT MRF funding. Specific activities are designed to enhance competitive funding for those scientists trained through and included in this project. Also, initial assessments of stakeholder needs will form the basis of a white paper that focuses attention on the needs of the scientific community and then project members will be encouraged to develop collaborative grants, with project activities designed to enhance the grant proposals.

It should be noted that new funding that spans research, education and training workshops will be one of the measures of project impacts. Furthermore, partnering on bioinformatics resources and educational initiatives will allow us to effectively leverage existing capacity that will be directed towards helping the animal industries.

Industry stakeholders have indicated the need for employees who have expertise in data management and analysis. We will partner with stakeholders to support and develop internships and similar experiential training opportunities so that those workforce needs are met.

A key component for sustaining capacity is to ensure that genomic resources are widely dispersed and easily accessible by multiple groups and users with different levels of expertise. With support from the EAB and industry stakeholders, we will develop policies for ensuring that the resources developed within this project are disseminated. For example, we foresee the distribution of analysis workflows/software to different bioinformatic platforms (e.g., CyVerse, Galaxy, GitHub, BioContainers).

## 2. Budget and Budget Narrative.

**OTT Multistate Research Funding requested:**

Salaries: We are requesting salary support to maintain the project website and email list (1 month p.a.); and summer stipends for students to develop and test bioinformatics analysis test cases (3 students each at 0.25 FTE, annually).

Fringes: Fringes are calculated at 43%.

Travel: We are requesting travel that will support EAB members attending annual meetings, training workshops, and community-stakeholder round table discussions. We also request funds to support students and early career investigators so they can attend bioinformatics/data science themed conferences. Travel costs include flights, accommodation, registration and meals.

Other: We are requesting the costs of open access publications for four reports on the outcomes of community and stakeholder discussions as well as funds for hosting the project server.

**Additional sources of funding:**

Project participants will be encouraged to seek additional funding to support the activities, data, resources and educational materials developed herein. NRSP-8 participants have been highly successful leveraging research funding using NRSP-8 tools and resources, with an average return on investment of more than $20 for each dollar provided to the NRSP-8 project. We intend to continue this incredible leveraged success by pursuing grant funding for conferences and meetings and education and training grants to expand and sustain the project's outreach activities. Moreover, by building stronger links with our industry stakeholders and including them in project management and decision making, we anticipate that we can obtain industry support that will offset the costs of meetings and training activities.

Figures for additional sources of funding are projected based on past support to NRSP-8 from industry partners and conference grants.

Salaries: Salary costs include support of industry research internships and university funding for summer stipends of students (1.5 students each at 0.25 FTE, annually).

Fringes: Fringes are calculated at 43%.

Travel: Travel costs include industry sponsorship of training workshops and stakeholder-focused discussions, as well as conference grants and fellowships that support students and early career investigators at bioinformatics/data science themed conferences. Travel costs include flights, accommodation, registration and meals.

# C. Integration and Documentation of Research Support:

## 1. Integration with Extension, academic, or international programs:

The proposed organizational structure is designed to ensure integration of capacity with Extension, academic and international programs. Current NRSP-8 membership is predominantly comprised of academicians with extension, teaching, and research appointments, thus promoting

communication through extension and outreach with stakeholders and industry representatives. We will build upon NRSP-8 sponsored conversations with industry representatives, which requires sustained conversations to develop knowledge, trust and understanding. Finally, it should be noted that the current membership of NRSP-8 includes 105 members from 49 institutions at 36 US states and six additional countries. A historic example of NRSP-8 supported international collaboration are efforts supporting the sequencing of key livestock genomes (10–14). More recent examples of NRSP-8 supported international collaborations include FAANG and AG2PI, both of which are high priority multi-species international projects that include and benefit NRSP-8 members. International partners will also be connected at scientific meetings (e.g., ASAS, ADSA, PAG, ISAG, SSR, WCGALP).

## 2. Engagement in project planning and implementation

The previous NRSP-8 program resulted in numerous new partnerships within the last five years. During this time, we held three stakeholder panels with representatives from all major USDA animal genome species. Industry participants represented both US and global animal industries, as well as animal genetics companies and breed associations. The feedback provided by these participants has been essential for developing the aims of the current proposal. During the same period, we created partnerships within the international FAANG and note that this project was initially proposed and developed as a direct consequence of the NRSP-8 project. Other partnerships that have developed from NRSP-8 activities include the AFRI funded projects developing pangenome resources for chicken, sheep, and cattle; several competitively funded telomere-to-telomere sequencing projects; the Agricultural Genome to Phenome Initiative (AG2PI), funded through the US Farm Bill; and strengthened linkages between NRSP-8 members and the AgBioData and NRSP-10 projects.

While NRSP-8 has been highly successful, ***this proposal is substantially different from the NRSP-8 project***. Funding for the NRSP-8 Species and Bioinformatics Coordinators is not requested in this project, and genomics tools and resources will not be the primary focus. Instead, the overall objective is the development of new genomics capacity across a multitude of areas. The project's implementation leverages lessons learned within NRSP-8. For example, we anticipate a smooth transition from NRSP-8 to the new leadership team by designating two active NRSP-8 members to serve as the first Co-chairs of the project and then electing chairs and coordinators from the new project's members. Stakeholders are engaged via the EAB and through individual activities that span the duration of the project. Members will be asked to identify their interest in specific aims and activities, and all members will be updated on all aims and activities at the annual meeting and via the annual written report. Moreover, all members are involved in selecting the NRSP's leadership team via the nomination and election processes.

## 3. Integration with multistate projects

This project will provide critical data and training resources to allow a wide range of scientists and researchers to use genomics and phenomics tools in their sub-disciplines of animal science. New workflows and case studies will be designed for connecting genomics with emerging technologies that are coming forward in animal agriculture. For example, the sensing and high-throughput phenotyping that validates behavioral responses in farm animals could be analyzed to identify the underlying genetics that control behavior. The project will also provide training through workshops associated with scientific meetings, websites and other virtual materials to the broader animal sciences community. Training will include sessions on a variety of topics such as how to use new analysis software, best practices when conducting genomic analyses, identifying phenomic

patterns and anomalies, and reusing (and archiving) previously collected data. It is anticipated that training may also be needed for new statistical approaches such as machine learning and AI. Obviously, these training modules will be useful to a wide range of discipline- specific multistate projects related to animal science and will also be pertinent in agricultural engineering and other related fields. Our plans for integration with other multi-state projects are outlined in the Business Plan.

# D. Outreach, Communications, and Assessment:

## 1. Communication Plan:

**Target Audience**. The primary beneficiaries of this NRSP include:

- Research scientists directly engaged in animal genetics, genome research, and animal breeding.
- Scientists utilizing publicly accessible genomic data in their respective animal science fields as well as complementary research in veterinary medicine, human physiology and medicine, bioinformatics, data science, biology, ecology, and rangeland science.
- Faculty, researchers, and graduate and postdoctoral students who want to apply new techniques in animal genomics to their research projects.
- Stakeholders, as defined and described in previous sections.

Support of research, education, and communication activities in animal genomics also benefits public consumers of US animal agriculture products. Outcomes include improvements in the quality, safety and wholesomeness of animal products, economic efficiencies, and environmental stewardship of animal production systems.

**Communication Plan**. The primary mechanism for distribution of NRSP results and accomplishments will be via annual reports publicly available on NIMSS (https://www.nimss.org/). However, we will also disseminate information about our activities and outcomes via peer-reviewed publications, white papers, and conference presentations. These products will be listed in the annual report posted on NIMSS and through other public channels, such as journal websites and relevant press releases. To engage with stakeholders and promote the project, these communications will also be featured on the project website, which will also contain training and educational resources developed within the project. The website will also be integrated into a list-serv that allows direct connection with members and stakeholders for disseminating information about upcoming project activities and opportunities.

## 2. Stakeholder engagement:

Our stakeholders are defined and described in the previous section, and details on how they will be integrated into this project are outlined in the Business Plan section. Briefly, stakeholders will be engaged at multiple points and will have an important role in shaping the outcomes of this project. Examples of stakeholder engagement are:

- A stakeholder advisory group (the Executive Advisory Board) will be created and meet annually with the project Leadership Team. The EAB will be critical in obtaining feedback on how project activities meet stakeholder needs.
- Stakeholders will be included on the project list-serve, which will be used to promote communication and dissemination of project activities and outcomes.

- Stakeholders will be encouraged to attend the annual project meetings.
- Workshops and training modules will be designed and focused on stakeholder needs.
- Surveys and discussions with stakeholders representing diverse industries and priorities are incorporated throughout the three Specific Aims.
- Experiential internships will link stakeholders with students, creating pipelines for data-savvy future employees as well as identify gaps in student knowledge.

## 3. Measuring accomplishments and outcomes:

The accomplishments of capacity-building projects are measured in terms of resources development, interdisciplinary collaborations, and knowledge transfer. The project Leadership Team will document accomplishments in annual progress reports, including:

- Published protocols and case studies that support bioinformatics analysis of animal genomics data sets and their integration into existing bioinformatics workflows.
- Development and use of guidelines to support best practice in bioinformatic analysis of genomics data sets and the reuse of data.
- Reports, white papers and publications describing outcomes from workshops, roundtables, and training events.
- Attendance of stakeholders at NRSP activities, including training workshops.
- Development of training, education and Extension materials and opportunities.
- Students participating in experiential education and internship opportunities.
- Publications and grant submissions that are outcomes of new collaborative opportunities established through this project.

Methods to measure program accomplishments have been described in previous sections of this proposal and include:

- An annual request from the Specific Aim Coordinators to all program participants to provide information on accomplishments and publications linked to the project.
- Routine searches of scholarly repositories, journal articles and professional reports.
- Compilations of project website use statistics.
- Participation in NRSP activities.

The impact of these accomplishments lies in their successful application to animal industries. This will be assessed by measuring usage of resources (e.g., online access and downloads), number of people who complete training and education activities, citations and downloads of publications and reports, stakeholder participation in the activities, and the ability of members to leverage project resources, activities and collaborations into research, education, extension and conference grants across federal and state agencies, as well as to research and philanthropic foundations.

## 4. Development of communication pieces:

The Leadership Team will organize annual reports which include accomplishments, outcomes and impacts. The report will be shared with NRSP members and stakeholders via the NIMSS reporting system and the project website and list serve. The Leadership Team and the External Advisory Board will identify a subset of project impacts and accomplishments that will have broader distribution across the Multistate Research Fund Impacts Program, State Agricultural Experiment Stations and leadership of the land-grant universities. As outlined above, project

accomplishments and outcomes will also be disseminated at scientific conferences and professional association meetings.

## 5. Data management plan:

***The goal of this capacity proposal is not to directly fund the acquisition of new data tools and resources but rather to support scientists who are using or wish to use them.*** NRSP members will be trained in best practices for data management, including analysis, sharing, and re-use of data. Members are expected to support responsible data management and routinely meet or exceed federal and community data management policy and best practices. This includes agreement to abide by the Toronto principles for data release, the Fort Lauderdale policy for rapid pre-publication release of data sets, and federal data sharing policies and requirements, including the USDA Guidelines for Data Management Planning. Goals of this project include the support of members in meeting these standards and to develop best practices for sharing data within the constraints of established commercial confidentiality.

Resources developed as a direct result of this project will be managed, archived, and made available to prospective users via publication in scientific journals, dissemination at meetings and conferences, through the project website and indirectly through links to other well-established platforms for resource sharing (e.g., GitHub, biocontainers, bioinformatic platforms and protocols.io).

## 6. Suggested mechanisms for distribution of the results of the research support project.

Many of the NRSP-8 coordinators and members have held advisory roles with USDA or other US agencies, offices, departments, or committees including:
- USDA ARS
- USDA APHIS
- USDA NIFA Southern Regional Aquaculture Center
- US DHS Analytic Exchange Program
- US DHS Cybersecurity and Infrastructure Security Agency
- US DHS Countering Weapons of Mass Destruction
- National Oceanic and Atmospheric Administration
- White House Rural Council.

As such, the NRSP-8 community expertise is well respected and communication channels to a diverse set of stakeholders are already in place. Furthermore, industry contacts serve in advisory roles for NRSP-8 and they provide and receive input that guides the research community. The NRSP-8 membership represents 49 institutions across 36 US states and six additional countries. Our established mailing list includes additional parties interested in this work, representing more than 3,000 individuals from 50 countries. These connections and networks will be used to distribute project results from the new project both nationally and internationally.

As described in previous sections, our goal is to have the project resources and reports widely disseminated, and this will ensure that the capacity developed by this project is sustainable after the funding is completed. While resources and accomplishments will be available on the project website, we expect that they will primarily be distributed to the community via well-established sharing platforms such as existing bioinformatics platforms (e.g., Galaxy, GitHub, Biocontainers, Data Carpentries and Protocols.io). These resources already have active communities and expertise for us to leverage and familiarity with these resources will also help our community to

develop additional expertise and collaborations. Resources developed as part of this NRSP will acknowledge NIFA Hatch funding, and members will be provided with reminder and suggested language for this. Resources and accomplishments will also be shared at the annual meetings, which will be associated with regular conferences and meetings to support outreach to our stakeholders.

**Literature Cited:**

1. Rexroad C, Vallet J, Matukumalli LK, Reecy J, Bickhart D, Blackburn H, et al. Genome to Phenome: Improving Animal Health, Production, and Well-Being – A New USDA Blueprint for Animal Genome Research 2018–2027. Frontiers in Genetics [Internet]. 2019 [cited 2022 Aug 17];10. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2019.00327

2. Mahfuz S, Mun HS, Dilawar MA, Yang CJ. Applications of Smart Technology as a Sustainable Strategy in Modern Swine Farming. Sustainability. 2022 Jan;14(5):2607.

3. Williams LR, Fox DR, Bishop-Hurley GJ, Swain DL. Use of radio frequency identification (RFID) technology to record grazing beef cattle water point use. Computers and Electronics in Agriculture. 2019 Jan 1;156:193–202.

4. Vinika G, West A, Bao Y, Brooks S, Staiger E. A Video Processing Pipeline for Equine Biomechanical Parameters Extraction and Gait Analysis. American Society of Agricultural and Biological Engineers. 2021;ASABE Annual International Virtual Meeting.

5. Amin S, Cuomo F, Kamal M. Comparative Analysis of Data Driven Prediction Modeling Strategies for Aquaculture Healthcare. In: 2021 International Conference on Innovative Computing (ICIC). 2021. p. 1–6.

6. Puma MJ. Resilience of the global food system. Nat Sustain. 2019 Apr;2(4):260–1.

7. U.S. Department of Agriculture Strategic Plan Fiscal Years 2022–2026.

8. Paltaki A, Michailidis A, Chatzitheodoridis F, Zaralis K, Loizou E. Bioeconomy and Livestock Production Nexus: A Bibliometric Network Analysis. Sustainability. 2021 Jan;13(22):12350.

9. Bashura JP. Food defense-"Back to the basics." Building the Future of Food Safety Technology. 2020;101–18.

10. Groenen MAM. A decade of pig genome sequencing: a window on pig domestication and evolution. Genetics Selection Evolution. 2016 Mar 29;48(1):23.

11. Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, et al. The sheep genome illuminates biology of the rumen and lipid metabolism. Science. 2014 Jun 6;344(6188):1168–73.

12. Tellam RL, Lemay DG, Van Tassell CP, Lewin HA, Worley KC, Elsik CG. Unlocking the bovine genome. BMC Genomics. 2009 Apr 24;10(1):193.

13. Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science. 2009 Nov 6;326(5954):865–7.

14. Warren WC, Hillier LW, Tomlinson C, Minx P, Kremitzki M, Graves T, et al. A New Chicken Genome Assembly Provides Insight into Avian Genome Structure. G3 Genes|Genomes|Genetics. 2017 Jan 1;7(1):109–17.